Oil discovery, political institutions and economic diversification

Article (Accepted Version)

Oil Discovery, Political Institutions and Economic Diversification[1]

Nouf Alsharif and Sambit Bhattacharyya[2]

7 August, 2018

Abstract: Diversification is touted as a desirable policy objective for oil rich nations because it reduces exposure to volatility. However, the empirical relationship between petroleum and diversification is not well understood. Here, we test the effect of giant oil discoveries on diversification using a panel dataset of 136 countries observed over the period 1962 to 2012. We notice non-oil sector export concentration 8 years after a discovery. However, we do not observe any effect on the structure of employment in non-resource and manufacturing sectors. Democratic political institutions moderate the export and employment concentration effects of petroleum discovery.

JEL classification: D72, O11

Key words: Oil Discovery; Political Institutions; Structural Change; Export Diversification

[1] We gratefully acknowledge comments by and discussions with Rabah Arezki, Matthias Busse, Kunal Sen, Gerhard Toews and the seminar/conference participants at Cambridge, Kiel, Siegen, Oxford, Sussex. All remaining errors are our own.

[2]Alsharif: Department of Economics, University of Sussex, email: N.Alsharif@sussex.ac.uk. Bhattacharyya: Department of Economics, University of Sussex, email: s.bhattacharyya@sussex.ac.uk.

1 Introduction

Export diversification and structural change is often touted as a desirable policy objective for petroleum rich nations. This is based on two related theoretical predictions. First, petroleum produces a highly concentrated structure of export revenue which is then exposed to volatility in petroleum prices. Such volatility in the long term is harmful for sustainable development and economic growth. Second, petroleum riches engender an economic structure that is highly concentrated and reliant on imported consumer goods and non-tradable services. Such structurally skewed economies are unable to deliver long term prosperity as the latter depends on rapid structural change away from non-tradables in the direction of tradables such as manufacturing and modern services.

In contrast, the classical trade theories of David Ricardo, Eli Heckscher and Bertil Ohlin predict that countries specialise and not diversify their exports and such pattern is largely dependent on the factor endowments rather than anything else. If a country is abundant in petroleum then it is perfectly predictable that it will have a petroleum dominated structure of exports.

Given such ambiguity between theoretical predictions and policy preferences, it is of enormous importance that policy advice be grounded in hard empirical facts. Yet, the empirical relationship between petroleum wealth, structural change, and export diversification is imperfectly understood. Is the relationship fundamentally driven by factor endowments? Do other factors such as politics and policy play a role? Given the common co-movement problem in observational data, to what extent we can attribute a causal relationship between petroleum wealth, structural change, and export diversification? What role political institutions and hence policy play in influencing the relationship between petroleum wealth and structural change?

A quick glance at the data in figures 1 and 2 suggest that petroleum wealth could

indeed be an important variable in understanding diversification patterns across countries and over time. Figure 1 presents scatter plots of the ratio of employment in the tradable and non-tradable sectors of an economy against oil dependence measured as a share of oil rent to GDP over the sample period 1962 to 2012. We can observe a clear positive pattern in panel A indicating non-tradable sector employment dominate over tradable sector employment in petroleum rich states. This pattern is somewhat weaker for countries with relatively strong political institutions (panel B) but stronger for countries with relatively weak political institutions (panel C). This pattern is repeated in figure 2 when we plot the ratio of non-resource exports to resource exports against petroleum wealth.

In this paper, we aim to systematically assess the role of petroleum in promoting or hindering structural change and export diversification. Using a panel dataset covering the period 1962 to 2012 and 136 countries we estimate the causal effect of giant oil discoveries on structural change and export diversification. All the estimates control for unobserved country-specific heterogeneity and time varying common shocks. We find evidence of non-oil export concentration 8 years after a giant oil discovery. The effect on the labour market structure however is absent. We find no effect of discovery on the structure of employment in the non-resource and manufacturing sectors.

The relationship between petroleum wealth and structural change could be influenced by the quality of political institutions and policy. Therefore we also test for potential heterogeneity in the relationship conditioned on institutional quality. We find that more democratic and inclusive political institutions moderate the concentration effects of petroleum discovery on exports but not the labour market. Similar trends are also observed with executive constraint as an alternative measure of institutions.

The paper makes the following original contributions. First, the paper presents an estimate of the causal effect of petroleum discovery on structural change and export

diversification. To the best of our knowledge, this result is entirely new. There are several recent studies that focuses on export diversification (see Imbs and Wacziarg, 2003; Cadot et al., 2011; and others) but none of them analyse the impact of petroleum discovery on diversification. Second, our dataset allows us to examine the effect of discovery news shock on tradable and non-tradable employment. This is a new result. Third, the paper also analyses the effect of institutional quality on diversification and presents new results.

Our identification strategy relies on the exogenous variation in the discovery dates of giant oil deposits. Our dataset codes a petroleum deposit (oil or/and gas including condensate) as giant if it has the capacity to generate 500 million barrels of ultimate recoverable oil or gas equivalent. Even though it is possible to identify the area where petroleum resources are likely to be found using geological data, it is not possible to accurately predict the timing of giant discoveries. Giant discoveries are rare and therefore the discovery dates of giant reserves are exogenous. One might argue that politicians and government could manipulate the announcement of the precise timing of discovery. Our data could potentially be immune to such possibility as the discovery dates are independently verified and recorded by multiple sources.

How exogenous is petroleum discovery? One could argue that petroleum discovery depends on exploration effort and effort depends on pre-existing political and economic conditions. It is also imperative that pre-existing economic and political conditions influence the structure of exports and the economy. Therefore, causal influence could run in the other direction from economic diversification to petroleum discovery. We test this argument empirically in section 2 and find that pre-existing economic and political variables do not predict discovery dates of giant petroleum reserves. Moreover, we also control for past petroleum discoveries as a proxy measure of discovery effort in all our estimated models. Finally, for the skeptics of the exogeneity of giant petroleum discoveries we estimate the

model using ‘out of region natural disasters’ and ‘oil reserves per capita’ as instruments for oil discovery.[3] These issues are discussed further in the remainder of the paper.

Our empirical strategy to estimate the heterogeneous effect of institutions rely on the Jones and Olken (2004) methodology. They observe institutional score prior to a random even and here we do the same by observing Polity2 score in the year prior to the discovery year. This mechanism allows the assessment of institutional quality one year prior to giant petroleum discovery so that the empirical estimate is not contaminated by any institutional change occurring simultaneously or thereafter.

Another related question is why we choose petroleum over other commodities. The reasons are as follows. First, the spatial distribution of petroleum discoveries are not skewed and therefore it offers the desired variation in the dataset to conduct a cross-national analysis (Wick and Bulte, 2009). If the discoveries were concentrated in one or two locations then any correlation between petroleum discoveries and diversification could be due to location specific unobservable confounding factors. Second, nearly most of recent resource discoveries have been in petroleum which makes the empirical analysis of its effect a worthwhile endeavor (Smith, 2015). Third, the resource curse literature attributes special properties to petroleum as a commodity which makes it important in the context of diversification and structural change. For instance, refining crude does not lead to a significant reduction in its weight. Therefore crude producing countries do not enjoy significant transport cost reducing advantages by setting up refineries locally. Furthermore, petroleum is capital intensive and has very little backward and forward linkages. In crude exporting poor countries they often remain as an enclave not contributing to the development of the rest of the economy (Alsharif et al., 2017).

Our paper is related to a literature on diversification. This literature mainly documents

[3] We follow Cotet and Tsui (2013) in using the abovementioned variables as instruments.

the pattern of diversification across countries and over time without exploring the role of petroleum wealth as a potential driver of such patterns. For example, Imbs and Wacziarg (2003) using sector level data on employment and production value added document a non-monotonic U-shaped pattern in diversification. In particular, Countries diversify at low levels of per capita income and up to a certain threshold. Beyond that threshold level of development countries experience sectoral concentration. This pattern is also confirmed by Cadot et al. (2011) when it comes to exports. Using a large database of 156 countries and tracking then over 19 years and 4991 product lines they find support for the U-shaped pattern in export diversification.

There is also a literature that examines the relationship between diversification patterns and growth. For example, Lederman and Maloney (2003), Hausmann et al. (2007) and Easterly at al. (2009) document that export patterns are path dependent and matter immensely for a nation state's long term growth prospect. Similar observations are made by a literature on structural change which documents large differences in labour productivity across traditional and modern sectors of an economy (McMillan and Rodrik, 2011; 2014; Rodrik, 2013). This literature argues that such differences in productivity is a major drag on the development potential of these economies. Again, none of these studies explore the role of petroleum discovery in hindering structural change.

The nearest to our paper is a literature on resource curse and Dutch disease. One of the key early theoretical contribution is from Corden and Neary (1982). Corden and Neary (1982) note that a resource windfall benefits the tradable primary export and non-tradable service sectors but at the expense of tradable non-resource (or manufacturing) sector. Subsequently, empirical research by Auty (2001), Gylfason (2001) and Sachs and Warner (2001, 2005) note that resource rich countries on average grow much slower than resource poor countries. This is further confirmed by studies that argue that natural resources may

lower the economic performance because they strengthen powerful groups, weaken legal frameworks, induce volatility, and foster rent-seeking activities (e.g., Ross, 2001; Ramey and Ramey, 1995; Koren and Tonreyro, 2007; and Besley, 2006). Others have argued whether natural resources are a curse or a blessing depends on country-specific circumstances especially institutional quality (eg., Mehlum et al., 2006; Bhattacharyya and Hodler, 2010, 2014; Bhattacharyya and Collier, 2014) and natural resource type (Isham et al., 2005).

In spite of the emerging consensus, the resource curse thesis is increasingly challenged by recent studies which uses more disaggregated spatial datasets and alternative identification strategies. For example, Smith (2015) finds no evidence of an oil curse using a cross-national dataset. Mamo et al. (2016) find very little evidence of a curse using spatial data on mineral discovery and nighttime lights in Africa. Similar observations are made for the US by Allcott and Keniston (2014) who examine the spatial effects of the Shale oil and gas boom at the county level.

The export diversification part of our paper closely relates to Bahar and Santos (2018). They argue that resource exports increase wages which makes several of the non-resource sectors expensive to operate. As a result non-resource exports become highly concentrated. They also show that a larger share of natural resources in exports lead to export concentration. They use commodity price shocks and oil discovery as instruments for the share of natural resource exports in their instrumental variable estimates. They also find that the non-resource export basket of resource rich countries are dominated by capital intensive products. Our results on non-resource export diversification in petroleum rich countries broadly concur with their results. However, unlike Bahar and Santos' (2018) instrumental variable approach we focus on the direct effects of oil discovery. Our empirical model allows us to examine the intertemporal effects of discoveries with lags and leads which yields new results. This is crucial as not all discoveries enter production and not all of them enter

production at the same time. On average it takes 8 to 10 years for a deposit to be developed and enter production. This allows us to distinguish between the expectation induced (after discovery but before the start of production) and revenue induced (after start of production) effects. These results are entirely new. In addition to examining export diversification, we are able to analyse diversification in employment, manufacturing employment, and manufacturing value added. In contrast, Bahar and Santos (2018) solely focus on export diversification.

The remainder of the paper is structured as follows: Section 2 analyses the effect of giant petroleum discovery on structural change and export diversification. In doing so, it carefully analyses the history of oil discovery, data, estimation strategy and results. Section 3 analyses the role of institutions and the heterogeneous effects of petroleum on tradable and non-tradable sectors. Section 4 presents additional robustness results and section 5 concludes.

2 Oil Discovery and Economic Diversification

2.1 What Drives Oil Discovery? Lessons from the History of Oil

Our identification strategy relies on the exogeneity of giant oil discovery. In other words, we argue that new giant oil discoveries are exogenous because they are independent of country specific factors. In this section, we explore the history of the oil industry and exploration across the globe. This exercise provides further credence to our thesis that oil discovery is orthogonal to country or market specific factors.

Yergin (1991) and Ross (2012) notes that Edwin Drake found oil in 1859 in Pennsylvania using the drilling method. Prior to Drake's drilling, oil was usually collected from water surface and used mainly for therapeutic purposes. Such use of oil dates back to 3000 B.C. Babylon and various other parts of the Middle East. Drake's invention of the drilling technology however significantly altered the oil landscape with the establishment of oil industry first in the United States, and later in the Russian Empire and some parts of East

Asia. With the outbreak of World War 1, the global demand for oil surged stimulating exploration effort. Increased exploration effort led to new discoveries and expansion of production all across the globe. Furthermore, new technologies such as the seismograph, aerial surface plotting, and micropaleontology significantly improved production in the 1920s and the 1930s. Following World War 2, production expanded to new locations in the French colonies of North and Sub-Saharan Africa (Algeria, Congo, Gabon, Libya, Nigeria). This expansion in production was driven by ever increasing military demand from imperial armed forces rather than anything specific to these countries. The post war period also witnessed expanded use of automobiles which further strengthened petroleum demand.

Introduction of more new technologies from the scientific disciplines of geochemistry, sedimentology, satellite imaging, and computing improved the prospects of new discovery even further. Offshore deep water drilling technologies also made discoveries feasible in locations which were inaccessible in the past.

The structure of the supply side however witnessed very few changes during this entire period even though petroleum production increased. The supply side of the oil market was dominated by seven major oil companies – the so called "seven sisters". These companies were Standard Oil of New Jersey (later Exxon), Standard Oil of California (later Chevron), Anglo-Iranian Oil Company (later BP), Mobil, Texaco, Gulf and Royal Dutch Shell. These companies controlled nearly the entirety of the oil market.

Therefore, what we learn from this historical detour is that giant petroleum discoveries are mainly driven by global factors, such as advancement in technology and increased demand. They appear to be unaffected by oil price changes in the 1970s (Smith, 2015). Smith (2015) notes that most of discoveries occurred before the 1970s price hike and prior to the oil shock prices were fairly low. They also appear to be exogenous to country specific factors. Nevertheless, we include country fixed effects in our model to account for

country specific factors. We also include time dummies to control for global factors such as technology and demand shocks.

2.2 Empirical Strategy and Data

We use a panel dataset covering up to 136 countries observed over the period 1962 to 2012.[4] To examine the effect of giant oil discovery on diversification we estimate the following model. The model is similar in spirit to Lei and Michaels (2014).

$$Div_{it+j} = \alpha_i + \omega_t + \gamma_1 Disc_{it} + \gamma_2 X_{it} + \varepsilon_{it} \quad (1)$$

where Div_{it+j} is the outcome variable (export diversification or structural change) in country i and year $t+j$, α_i is a country dummy variable accounting for country fixed effects, ω_t is a year dummy variable controlling for time varying common shocks, $Disc_{it}$ is an indicator of a giant oilfield discovery in country i and year t, and X_{it} is the number of years with resource discoveries in the last ten years (from $t-10$ to $t-1$). We estimate this model for different leads j, where in most cases $j \in \{2,4,6,8,10\}$. This is important for the purpose of tracking long terms effects of the oil discovery shock. In order to check robustness of the coefficient estimate of interest, we include additional covariates in the extended version of this specification. The additional covariates include GDP per capita and GDP per capita squared.

Our main coefficient of interest here is γ_1, which is the effect of an oil discovery shock on diversification. If a giant oil discovery shock leads to export and structural concentration then we would expect γ_1 to be positive and statistically significant. Any indication otherwise would indicate that this is not the case.

Our main dependent variables are the export diversification and structure change measures. We use sectoral data on non-resource employment and exports to compute concentration indices. Employment data is sourced from the ILO and UNIDO whereas the

[4] Due to data limitations, most but not all specifications cover 136 countries. In most specifications, the panel is unbalanced. Appendix A1 presents a list of countries included in the sample.

export data is sourced from WITS. We also use sectoral value added data for manufacturing from UNIDO to check the robustness of our results. The number of countries in the dataset is up to 136 with varying degree of development. The observations are annual, covering the period from 1962 to 2012. We compute several measures of diversification and most of them are inspired by the income equality literature. Our preferred measure of diversification is the Gini coefficient. Nevertheless, we also estimate our model using other indices such as Theil and Herfindahl-Hirschman (HH) and the regression results are similar.[5] This is not surprising given that all the indices are highly correlated. Tables 1 and 2 presents descriptive statistics and correlation coefficients of these indices respectively. Appendix A3 also provides further details on how the indices are computed, the underlying data structure and source. The Gini coefficient varies between 0 and 1 and a higher Gini would imply that exports or the labour market is highly concentrated whereas a lower Gini would signal diversification.

There is no consensus in the literature with regards to the most appropriate measure of diversification. Imbs and Wacziarg (2003) for example report Gini, HH index and the Coefficient of Variation. In contrast, Cadot et al. (2011) drop the coefficient of variation and only use Gini and the HH index. McMillan and Rodrik (2011) only focus on the Coefficient of Variation while ignoring Gini and the HH index. We run our regressions using all three indices (Gini, Theil and HH) but we only report the Gini index in the main paper. All other results involving HH index and Theil index are reported in the Online Appendix.

According to our data from the ILO, Algeria in 1984 appears to be the most diversified in terms of employment in the non-resource sector. Greece in 2006 appears to have highly diversified exports whereas exports in Libya over the period 1976 to 1981 appears to be highly concentrated.

Oil discovery data is sourced from Lei and Michaels (2014), which is based on a

[5] These results are reported in the Online Appendix.

dataset by Horn (2004). Horn reports the date of discovery, the name of the discovering country, and a number of other variables, for 910 giant oilfields discovered both onshore and offshore over the period 1868 to 2003. As we have mentioned earlier, to qualify as a giant discovery, an oilfield must contain ultimate recoverable reserves of at least 500 million barrels of oil equivalent.

We plot the number of giant oil discoveries over time in figure 3. The plot shows that discoveries peaked in the 1960s and 1970s, while declining significantly in the 1980s. Double-digit discoveries returned in the late 1990s but the declining trend continued in the noughties. Of the total 910 giant oilfield discoveries covered by Horn (2004), only 364 are used in this paper which took place within our sample period 1962 to 2003.[6] The diversification data that we have from UNIDO, ILO and WITS runs till 2012 and therefore giving us the opportunity to analyse the effect of a discovery shock up to a decade later.

The discovery episodes (364 country-year observations) are 5.2 percent of the total sample size and therefore are rare events. This is further confirmed in table 3. We observe that 40 percent of the giant discoveries during our sample period came from Asia followed by Europe (19 percent), Africa (17 percent), South America (10 percent), North America (9 percent) and Oceania (5 percent). The treatment group consists of 64 countries who experienced at least one giant oil discovery during the sample period. The control group consists of 72 countries who have never experienced any giant oil discoveries. This provides an opportunity for a balanced comparison.

Finally, we also use GDP per capita as a control variable and these figures are sourced from the World Bank's World Development Indicators.

Before we engage in estimating the average effect it is probably worthwhile analysing

[6] The aim here is to estimate the intertemporal effect of the same discovery shocks. Using different discovery shocks for different lag length does not make the time lapse effect of these shocks comparable. Nevertheless, we also estimate our main regression using different discovery shocks for different time lags and the result is robust.

some country specific trends. In figure 4 we examine the effect of giant oil discoveries on the structure of exports. We deliberately choose countries with very different political institutions. We observe export concentration post giant oil discovery news shock but the effect appears to be somewhat mild for democracies (Denmark and Spain) as opposed to non-democracies (Egypt and the Republic of Congo). The discoveries displayed in these figures are not necessarily exclusive; there might be more giant discoveries in other years.

Figure 5 focuses on the industrial (manufacturing) employment in Egypt, Indonesia, Norway, and Australia. Irrespective of the institutional background, it appears that countries experience concentration in industrial employment post petroleum discovery.

2.3 Evidence

2.3.1 Identification

Our underlying identification assumption is that giant petroleum discoveries in a country are exogenously timed and are orthogonal to the underlying economic conditions of that country. Therefore, before we start testing the impact of giant petroleum discoveries on diversification, it is worthwhile testing the underlying identification assumption.[7] To do that, we estimate a fixed-effects logit model in table 4, where the independent variables are lags of diversification in non-resource and manufacturing sectors and other political economic variables (lagged polity2 score, lagged GDP growth, lagged GDP per capita growth, lagged government expenditure growth, and lagged investment growth) and the dependent variable is a dummy variable equal to one in the year of a giant petroleum discovery. If the identification assumption is invalid then we would observe past changes in political and economic variables would predict the petroleum discovery dates. As expected, we find that the key variable of interest – diversification – as well as changes in other economic and political variables do not predict giant oil discoveries. We also estimate the model with a lag

[7] Note that Smith (2015) also uses a similar test and finds similar results.

length of 2 and the result remains unaffected.

2.3.2 Oil Discovery and Diversification: Baseline Results

In table 5 we estimate the effect of petroleum discovery on non-resource export diversification and structural change. In panel A we notice export concentration 8 years post giant oil discovery and the effect is statistically significant. The magnitude of the non-resource export concentration effect somewhat declines after a decade post discovery but still remains significant. This is in line with the expectation that it takes 5-6 years post discovery for reserves to come into production and hence we notice a delayed concentration effect. The magnitude of the coefficient is one tenth of the sample standard deviation, a relatively small effect. To put the magnitude of the coefficients in perspective, let us consider Angola in 1990. It has a non-resource export diversification Gini of 0.8765 very close to the sample maximum of 0.8888. A giant oil discovery in 1990 will increase non-resource export concentration Gini in Angola to 0.8895 in 1998 and to 0.9005 in 2000. Note that Nigeria registers very high concentration of non-resource exports with Gini of 0.8259 in 2003. Our estimates predict that a giant oil discovery in Angola will push it significantly higher than Nigeria in terms of export concentration. Panels B and C deals with non-resource employment and manufacturing employment. The distribution of employment in both non-resource and manufacturing sectors appear to be unaffected by the discovery news shock. This is not entirely surprising. Modern petroleum industry is extremely capital intensive and therefore petroleum discovery shocks are not expected to affect the labour market in a major way.

Lei and Michaels (2014) point out that petroleum discoveries in a country's recent past could raise the likelihood of additional discoveries in the immediate future. It could also significantly reduce the likelihood of a giant discovery if the country has low potential in terms of reserves. Therefore, all specifications reported in table 5 controls for giant petroleum

discoveries in the last ten years. They also control for country fixed effects and time varying common shocks (year dummies). Figure 6 plots the effect 4 years before (t-4) and 10 years after (t+10) the discovery shock.

2.3.3 Do the Size of Petroleum Discovery Matter?

So far we have concentrated on giant oil discoveries. What if smaller petroleum discoveries affect diversification disproportionately more than giant discoveries? Smaller discovery shocks however are unlikely to be exogenous. Nevertheless, in table 6 we examine the effect of non-giant discovery shocks on non-resource export diversification and structural change. The data for non-giant oil discoveries is sourced from Horn (2004) which classifies oil discoveries into giant and major. Major deposits contain ultimate recoverable reserves of less than 500 million barrels of oil equivalent. We find evidence of export concentration but no effect on the structure of employment in the non-resource and manufacturing sectors.

Giant oilfield discoveries themselves vary in terms of sizes and could be a potential source of heterogeneity. Therefore, one could question the justification of lumping all these discoveries of varying degree and size under one group namely giant. In other words, what if the size distribution among the giant discoveries matter for diversification? To investigate, we test the relationship between the different sizes of giant discoveries and diversification. In table 7, we divide the giant oilfield discoveries normalized by population by their respective size. In particular, we divide them into four quartiles based on the size of the estimated ultimate recoverable reserves divided by population. We notice that the effect is small for quartiles 1 to 3 and mostly insignificant. Some coefficients are significant but only at the 10% level. The strongest effect is registered by the largest discoveries in quartile 4. The non-resource export concentration effect is strong and statistically significant after 8 years. This result supports the view that the super-giant discoveries wield the most influence on diversification and structural change. Note that quartiles calculated using raw discovery size

also has the same effect.

3 Oil Discovery, Political Institutions and Diversification

3.1 Empirical Strategy and Data

To examine the effect of political institutions we estimate the following modification of equation (1).

$$Div_{it+j} = \theta_i + \lambda_t + \delta_1 Disc_{it} + \delta_2 Disc_{it} \times INS_{it-1} + \delta_3 X_{it} + \nu_{it} \quad (2)$$

where INS_{it-1} is a measure of the quality of political institutions in country i and year $t-1$. The resource curse literature emphasize the role of political institutions in influencing the relationship between natural resources and economic development. Therefore it is worthwhile testing whether political institutions also affect the relationship between petroleum discovery and diversification.

We use polity 2 score from the Polity IV database as a proxy measure of political institutions or democracy. Ross (2001) documents that measures of political institutions or democracy could be endogenous to petroleum wealth. Hence we use lagged polity 2 score to account for the quality of political institutions before petroleum discovery. The variable varies between -10 and +10 with a higher score indicating better quality political institutions. The advantages of using the polity 2 variable is that it covers a broad cross-section of countries throughout our sample period. It is also conceptually attractive given that it codes formal constraints that are placed on the executive. Nevertheless, we also use executive constraint as an alternative measure of political institutions in the robustness section and the results are robust.

We are interest in the partial effects of a petroleum discovery shock and hence the coefficients δ_1 and δ_2. If a giant petroleum discovery shock leads to export and structural concentration in the non-resource sector then we would expect δ_1 to be positive and

statistically significant. If better quality political institutions moderate that effect then we would expect δ_2 to be positive and significant.

3.2 Evidence

In table 8 we report coefficient estimates of δ_1 and δ_2. In panel A we find strong evidence of export concentration 8 years after a giant petroleum discovery. This concentration effect stays statistically significant 10 years after a discovery. The concentration effect is moderated by better quality political institutions as the coefficient on the interaction term is negative and significant 8 years after discovery.

In panels B and C we check the effect of discovery on the internal structure of the economy and especially the labour market in the non-resource and manufacturing sectors. We do not find any statistically significant effect of discovery on employment.

4 Robustness

The quality of political institutions are dependent on the constraints that are imposed on the chief executive. Therefore, it is important to analyse the direct effect of such constraints. In table 9 we replace the polity 2 variable with executive constraints. In panel A we find that the export concentration effect of discovery remains unaffected. This effect is moderated by executive constraints only 8 years after the giant discovery. In panel B we notice strong concentration effect on employment in the non-resource sector 6 years post discovery and beyond. This effect is also moderated by higher levels of executive constraint. Panel C deals with manufacturing employment only with data from the UNIDO. We do not observe any statistically significant effect of giant petroleum discovery.

An alternative identification strategy is to use oil reserves and natural disasters as instruments for giant petroleum discoveries. In table 10 we follow Cotet and Tsui (2013) and use oil reserves and natural disasters as instruments for oil discoveries. The oil reserves instrument is log of oil reserves calculated for each country-year by subtracting cumulative

production from cumulative discovery. The data is sourced from the Association for the Study of Peak Oil (ASPO). Known oil reserves are likely to positively influence future oil discoveries. The other instrument is the log of out-of-region natural disaster where five kind of disasters are considered: earthquake, volcano, mass movement, storm and flood. Cotet and Tsui (2013) describes out-of-region disasters as the value of all disaster damages minus the value of own region damages. Natural disasters could affect oil discoveries through unexpected interruptions in exploration work. Therefore, it can serve as an instrument for discovery. However, natural disasters could also affect diversification through other channels therefore may not be ideally suited to satisfy the exclusion restriction. In panel A we find evidence of concentration 10 years after discovery. In panels B and C we also find evidence of concentration in employment in the non-resource and manufacturing sectors. This concentration effect is moderated by better quality political institutions.

Arezki et al. (2017) argue that large oil discoveries are not exogenous as countries with open economies are more likely to make discoveries. Therefore, we run table 5 panel A with the “openness” dummy variable as an additional control variable. The results do not change.

Finally, we also re-estimate tables 5 and 8 using alternative measures of diversification. In particular, instead of using Gini coefficient we use Theil index and Herfindahl-Hirschman (HH) index as measures of diversification and our main result remains unchanged. These results are reported in the Online Appendix. Furthermore, we also calculate Gini using 2 digit sectors and re-estimate panel A of table 5. Results are robust.

5 Conclusions

Using petroleum resources to promote a diversified economy has been a challenge especially for petroleum rich developing countries. This is in addition to the challenges faced by these countries in terms of capital constraints, attracting private investments into the petroleum

sector, and maintaining a disciplined fiscal regime to capture revenue and reduce macroeconomic volatility (Venables, 2016). Both national and international policy circles acknowledge these challenges and promote economic diversification as a desirable objective. In spite of such rare policy consensus, our knowledge of the empirical relationship between petroleum wealth and diversification is largely incomplete. Causal direction of the relationship is also partially understood.

In this paper, we systematically assess the role of petroleum wealth on diversification. In particular, we estimate the causal effect of giant oil discoveries on structural change and export diversification. We find evidence of non-oil export concentration. This export concentration effect is somewhat moderated by better quality political institutions. We also find that countries experience concentration of employment in the non-tradable sector relative to the tradable sector two years post petroleum discovery but the effect is short lived.

We contribute to the literature by estimating the effect of oil discoveries on export diversification and structural change. This is a new result. Our dataset allows us to distinguish between tradable and non-tradable sector employment which is new. We also introduce political institutions into this literature.

The diversification challenge for petroleum rich economies is not exclusively a developing country problem. In fact a quick look at the export composition data reveals that even for a developed nation such as Norway with good political institutions resource exports have reached almost 50% of total exports in 2013 crowding out other tradables. In fact Norway's share of manufacturing exports dropped from approximately 70 percent in 1972 to only 17 percent in 2013. The shares per se should not be a concern, but the association with less non-resource output is worthwhile noting. Our regression analysis also confirms this trend. This underscores the strength of the specialization argument put forward by the classical trade theory literature of Ricardo, Heckscher and Ohlin.

Appendices

A1. Countries and the Type of Government at the Time of Discovery:

Democratic	**Autocratic**
Argentina	Afghanistan
Australia	Albania
Bolivia	Algeria
Brazil	Argentina
Canada	Azerbaijan
Colombia	Bangladesh
Congo, Rep	Brazil
Denmark	Cameron
Ecuador	China
France	Colombia
India	Congo, Rep
Indonesia	Cote d'Ivoire
Iran	Egypt
Italy	Equatorial Guinea
Malaysia	Gabon
Mexico	Hungary
Netherlands	Indonesia
New Zealand	Iran
Nigeria	Iraq
Norway	Kazakhstan
Pakistan	Kuwait
Papua New Guinea	Libya
Peru	Mexico
Philippines	Morocco
Romania	Myanmar
Russia	Nigeria
Spain	Oman
Thailand	Qatar
Trinidad & Tobago	Saudi Arabia
United Kingdom	Sudan
United States	Thailand
Venezuela	Tunisia
	Turkmenistan
	USSR
	United Arab Emirates
	Venezuela
	Vietnam

A2. Countries and the Type of Government One Year Prior to Discovery

Country	Discovery year	Type of government (polity2)	Country	Discovery year	Type of government (polity2)
Argentina	1971	Autocratic (-9)	Mexico	1951	Autocratic (-6)
	1977	Autocratic (-9)		1952	Autocratic (-6)
	1989	Democratic (8)		1958	Autocratic (-6)
	1996	Democratic (7)		1966	Autocratic (-6)
Brazil	1965	Autocratic (-3)		1972	Autocratic (-6)
	1968	Autocratic (-8)		1975	Autocratic (-6)
	1972	Autocratic (-9)		1976	Autocratic (-6)
	1984	Autocratic (-3)		1977	Autocratic (-6)
	1985	Autocratic (-3)		1979	Autocratic (-3)
	1987	Democratic (7)		1980	Autocratic (-3)
	1989	Democratic (7)		1982	Autocratic (-3)
	1993	Democratic (8)		1990	Autocratic (0)
	1996	Democratic (8)		1998	Democratic (6)
	1999	Democratic (8)	Nigeria	1958	NA
	2001	Democratic (8)		1959	NA
	2002	Democratic (8)		1962	Democratic (8)
	2003	Democratic (8)		1963	Democratic (8)
Colombia	1956	Autocratic (-5)		1964	Democratic (8)
	1973	Democratic (7)		1965	Democratic (7)
	1992	Democratic (9)		1967	Autocratic (-7)
	1993	Democratic (9)		1968	Autocratic (-7)
Congo, Rep	1969	Autocratic (-7)		1970	Autocratic (-7)
	1971	Autocratic (-7)		1973	Autocratic (-7)
	1983	Autocratic (-8)		1981	Democratic (7)
	1995	Democratic (5)		1989	Autocratic (-7)
Indonesia	1969	Autocratic (-7)		1990	Autocratic (-5)
	1970	Autocratic (-7)		1996	Autocratic (-6)
	1971	Autocratic (-7)		1998	Autocratic (-7)
	1972	Autocratic (-7)		1999	Autocratic (-1)
	1973	Autocratic (-7)		2000	Democratic (4)
	1974	Autocratic (-7)		2001	Democratic (4)
	1982	Autocratic (-7)		2002	Democratic (4)
	1991	Autocratic (-7)	Thailand	1973	Autocratic (-7)
	1994	Autocratic (-7)		1980	Democratic (2)
	1995	Autocratic (-7)		1995	Democratic (9)
	1996	Autocratic (-7)	Venezuela	1954	Autocratic (-3)
	1997	Autocratic (-7)		1955	Autocratic (-3)
	1999	Autocratic (-5)		1957	Autocratic (-3)
	2000	Democratic (6)		1958	Autocratic (-3)
Iran	1958	Autocratic (-10)		1979	Democratic (9)
	1960	Autocratic (-10)		1980	Democratic (9)
	1961	Autocratic (-10)		1986	Democratic (9)
	1962	Autocratic (-10)		1988	Democratic (9)
	1963	Autocratic (-10)		1999	Democratic (8)
	1964	Autocratic (-10)		2002	Democratic (6)
	1965	Autocratic (-10)			
	1966	Autocratic (-10)			
	1967	Autocratic (-10)			
	1968	Autocratic (-10)			
	1969	Autocratic (-10)			
	1972	Autocratic (-10)			
	1973	Autocratic (-10)			
	1974	Autocratic (-10)			
	1975	Autocratic (-10)			
	1976	Autocratic (-10)			
	1978	Autocratic (-10)			
	1980	Autocratic (0)			
	1988	Autocratic (-6)			
	1991	Autocratic (-6)			
	1992	Autocratic (-6)			
	1993	Autocratic (-6)			
	1994	Autocratic (-6)			
	1995	Autocratic (-6)			
	1999	Democratic (3)			
	2000	Democratic (3)			
	2001	Democratic (3)			

Note: This table only lists countries that experienced at least one regime switch.

A3. Data Appendix:

Employment data

Sectoral employment data are from International Labor Office (ILO, 2013) and United Nations Industrial Development Organization (UNIDO, 2012). ILO data covers 127 countries, while UNIDO covers 125 countries. The ILO data includes all economic activities at the 1-digit level between 1969 and 2008. Sectoral shares are in percentages. The unbalanced panel has 2369 observations (country-year). The ILO dataset reports employment in different classifications: some countries use the ISIC-revision 2, others moved to ISIC-revisions 3 and 4 in recent years, and some are using their own national classification. Employment data in the more disaggregated ISICrev3 and ISICrev4 were aggregated to ISICrev2, following Imbs and Wacziarg (2003), Timmer and de Vries (2008) and McMillan and Rodrik (2011). If a country reports two revisions, the lower one is used. Official estimates are preferred over labor surveys. Data not following ISIC conventions are dropped. Table B1 shows the concordance between ISICrev3 and ISICrev2.

Table B1: different classifications between ISIC revisions 2 and 3*

ISIC-Revision 2	ISIC-Revision 3 Equivalent
1. Agriculture, Hunting, Forestry and Fishing	A. Agriculture, Hunting and Forestry B. Fishing
6. Wholesale and Retail Trade and Restaurants and Hotels	G. Wholesale and Retail Trade; Repair of Motor Vehicles, Motorcycles and Personal and Household Goods H. Hotels and Restaurants
8. Financing, Insurance, Real Estate and Business Services	J. Financial Intermediation K. Real Estate, Renting and Business Activities
9. Community, Social and Personal Services	L. Public Administration and Defense; Compulsory Social Security M. Education N. Health and Social Work O. Other Community, Social and Personal Service Activities P. Households with Employed Persons

* McMillan and Rodrik (2011) and Timmer and de Vries (2008)

ILO data sometimes have sudden big changes in numbers in certain sectors, as countries sometimes change their calculation method even if the same classification/revision is used. This is taken into consideration in this study, by dropping the observations that reports these sudden changes making the panel more harmonized.

Our alternative data source is UNIDO, which covers manufacturing activities only at the 3-digit level

of disaggregation (the main 23 industrial sectors) between 1963 and 2010 (INDSTAT2). (INDSTAT4 disaggregates to 4-digit level but only goes back to 1985). The UNIDO dataset is consistent over the years and did not need adjustment. The unbalanced panel has 3564 employment observations (country-year).

Exports data

Exports data are from the World Integrated Trade Solution (WITS), which is collaboration between the World Bank and the United Nations Conference of Trade and Development (UNCTAD). The export data covers 133 countries. Data is selected in SITC-1-digit aggregation containing the main 10 trade sectors (0 Food and live animals; 1 Beverages and tobacco; 2 Crude materials, inedible, except fuels; 3 Mineral fuels, lubricants and related materials; 4 Animal and vegetable oils, fats and waxes; 5 Chemicals and related products, not elsewhere specified; 6 Manufactured goods classified chiefly by material; 7 Machinery and transport equipment; 8 Miscellaneous manufactured articles; 9 Commodities and transactions not classified elsewhere in SITC. Values are reported in constant 1000 USD with base year 2000. The unbalanced panel has 4575 observations (country-year). The WITS data values are consistent over the years and did not need any adjustment.

Diversification Indicators

Computation of these measures is done in Stata.[8]

Table B2: The main differences between the chosen concentration measures[9].

Index	Distance Concept	Decomposable?	Independence of input scale & population size?	Range in interval [0,1]?
Gini	Depends on rank ordering	No	Yes	Yes
Theil	Proportional	Yes	Yes	No
HHI	Absolute differences	Yes	No: decreases with population	Yes: but min>0

We calculate diversity for all sectors, and for all non-resource sectors. Specifically, in the ILO data we exclude “Mining and Quarrying”, and in the WITS exports data we exclude “Crude material, inedible, except fuels”, “Mineral Fuels, lubricants and related materials” and “Commodities not classified according to kind”. The UNIDO data does not cover resource sectors at all.

[8]See AINEQUAL: Stata module to compute measures of inequality.
[9]See Cowell (2011).

References

Allcott, H. and D. Keniston (2014). Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America. Unpublished Manuscript.

Alsharif, N., S. Bhattacharyya, and M. Intartaglia (2017), "Economic Diversification in Resource Rich Countries: History, State of Knowledge and Research Agenda" *Resources Policy*, 52, 154-164.

Arezki, R., F. van der Ploeg, and F. Toscani (2017), "The Shifting Natural Wealth of Nations: The Role of Market Orientation," Unpublished Manuscript.

Auty, R. (2001). The political economy of resource-driven growth. *European Economic Review*, 5(4):839–846.

Bahar, D., and M. Santos (2018), "One more resource curse: Dutch disease and export concentration," *Journal of Development Economics*, forthcoming.

Besley, T. (2006). Principled Agents? Princeton University Press, Princeton NJ.

Bhattacharyya, S. and P. Collier (2014), "Public Capital in Resource Rich Economies: Is there a Curse?" *Oxford Economic Papers*, 66 (1), 1-24.

Bhattacharyya, S., and R. Hodler (2010). Natural Resources, Democracy and Corruption, *European Economic Review*, 54, 608-621.

Bhattacharyya, S., and R. Hodler (2014). Do Natural Resource Revenues Hinder Financial Development? The Role of Political Institutions, *World Development*, 57, 101-113.

Cadot, O., Carrère, C., & Strauss-Kahn, V. (2011). Export diversification: What's behind the hump? *Review of Economics and Statistics*, *93*(2), 590-605.

Corden, W. M., & Neary, J. P. (1982). Booming sector and de-industrialisation in a small open economy. *The economic journal*, 825-848.

Cotet, A. and K. Tsui (2013). "Oil and Conflict: What Does the Cross Country Evidence Really Show?" *American Economic Journal: Macroeconomics*, 5(1): 49–80.

Cowell, F. (2011). *Measuring inequality*: Oxford University Press.

Easterly, W., Reshef, A., & Schwenkenberg, J. M. (2009). The power of exports. *World Bank Policy Research Working Paper Series, Vol*.

Gylfason, T. (2001). Natural resources, education, and economic development. *European Economic Review*, 45:847–859.

Hausmann, R., J. Hwang, and D. Rodrik (2007). What You Export Matters. *Journal of Economic Growth*, 12:1–25.

Horn, M. (2004). Giant Fields 1868-2004 (CD-ROM), Houston.

Imbs, J., & Wacziarg, R. (2003). Stages of diversification. *American Economic Review*, 93(1): 63-86.

Isham, J., L. Pritchett, M. Woolcock, and G. Busby (2005). The varieties of resource experience: natural resource export structures and the political economy of economic growth. *World Bank Economic Review*, 19, 141-174.

Jones, B. F., & Olken, B. A. (2004). Do leaders matter? National leadership and growth since World War II. *National Leadership and Growth Since World War II (March 2004).*

Koren, M., & Tenreyro, S. (2007). Volatility and development. *The Quarterly Journal of Economics*, 243-287.

Lederman, D., & Maloney, W. F. (2003). Trade structure and growth. *World Bank Policy Research Working Paper*(3025).

Lei, Y-H., and G. Michaels (2014). Do Giant Oilfield Discoveries Fuel Armed Conflicts? *Journal of Development Economics,* 110: 139–157.

Mamo, N., R. Arezki, S. Bhattacharyya, and A. Moradi (2015). Economic Consequences of Mineral Discovery and Extraction in Sub-Saharan Africa: Is There a Curse? Unpublished manuscript, University of Sussex.

McMillan, M. S., & Rodrik, D. (2011). Globalization, structural change and productivity growth: National Bureau of Economic Research.

McMillan, M. S., & Rodrik, D. (2014). Globalization, Structural Change and Productivity Growth, with an Update on Africa. *World Development*, 63: 11-32.

Ramey, G. and V. Ramey (1995), "Cross-country Evidence on the Link Between Volatility and Growth," *American Economic Review* 85(5): 1138-51.

Rodrik, D. (2013). Unconditional Convergence in Manufacturing. *Quarterly Journal of Economics*, 128(1):165–204.

Ross, M. (2001). Does oil hinder democracy? *World Politics, 53*(03), 325-361.

Ross, M. (2012). *The oil curse: how petroleum wealth shapes the development of nations*: Princeton University Press.

Sachs, J., and A. Warner (2001), "The curse of natural resources," *European Economic Review*, 45, 827-838.

Sachs, J., and A. Warner (2005), "Natural Resource Abundance and Economic Growth," in Meier, G. and J. Rauch (eds.) *Leading Issues in Economic Development*, OUP: New York.

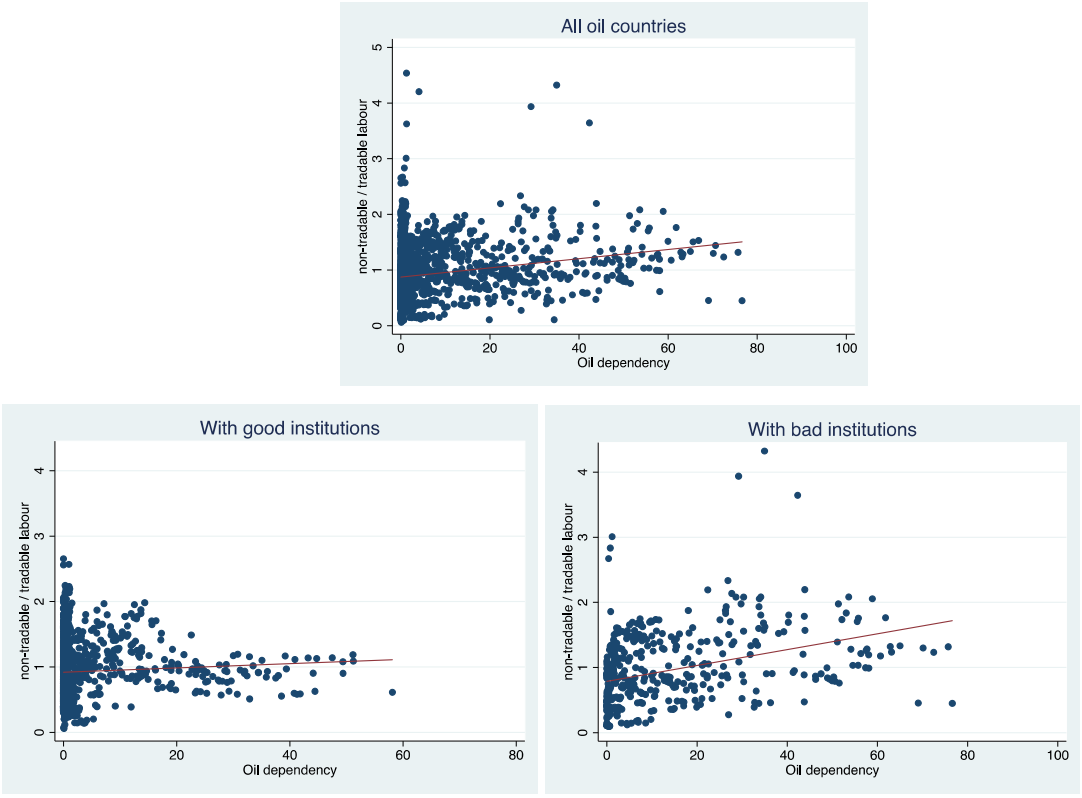
Smith, B. (2015). The resource curse exorcised: Evidence from a panel of countries. *Journal of Development Economics, 116*, 57-73.

Timmer, M. P., & de Vries, G. J. (2008). Structural change and growth accelerations in Asia and Latin America: a new sectoral data set. *Cliometrica, 3*(2), 165-190.

Venables, A. (2016), "Using natural Resources for Development: Why Has It Proven So Difficult?" *Journal of Economic Perspectives*, 30(1): 161-184.

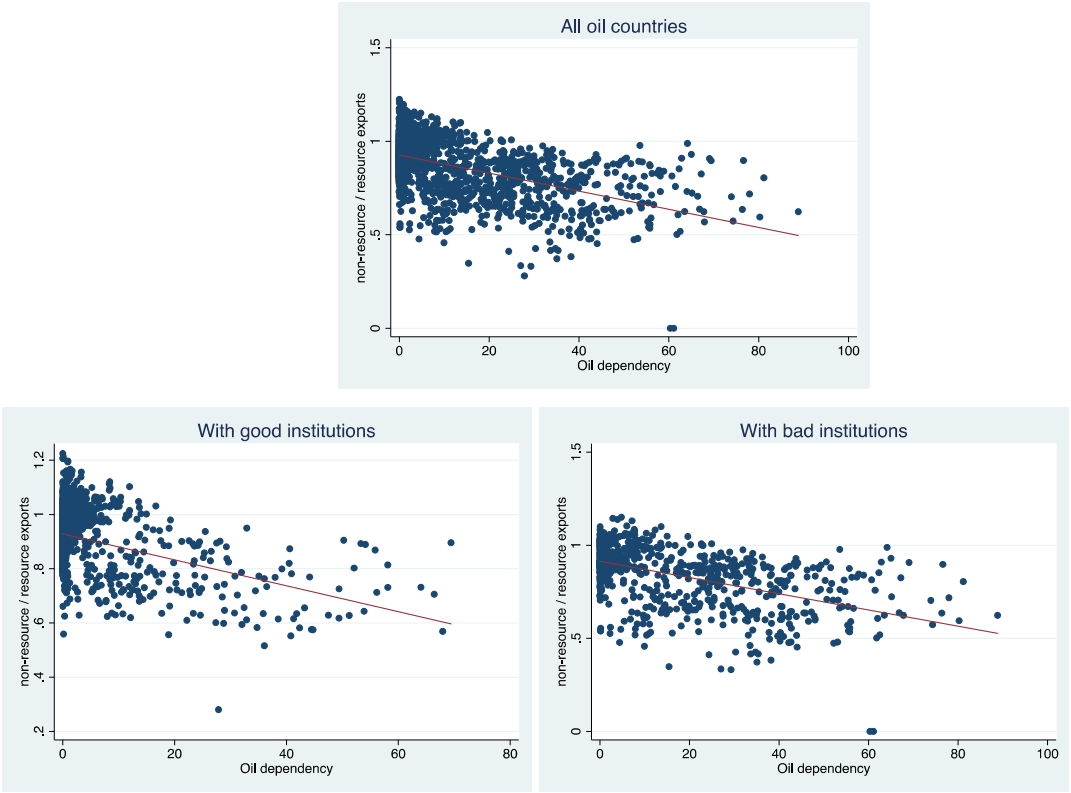
Yergin, D. (1991). *The Prize: The Epic Quest for Oil, Money and Power*, Simon & Schuster: New York.

Figure 1: Oil, the Dutch Disease and institutions (a) all oil countries, (b) with good institutions (c) with bad institutions



Notes: Resource movement effect; as labour move from tradable to non-tradable sectors with higher oil abundance. X-axis is oil dependency measures by oil rent share in GDP, data from the World Bank. Y-axis is the relative employment share in non-tradable to tradable sectors within the ILO data. Panel (a) includes all countries in our dataset. Countries in panel (b) are: Australia, Austria, Canada, Columbia, Czech Republic, Denmark, France, Georgia, Germany, Greece, Israel, Italy, Japan, Lithuania, Malaysia, Netherlands, Norway, Serbia, Slovak Republic, Slovenia, Trinidad and Tobago, Turkey, Ukraine, United Kingdom, United States of America, Venezuela. Countries in panel (c) are: Algeria, Angola, Azerbaijan, Bahrain, Belarus, Cameroon, China, Cuba, Egypt, Gabon, Iran, Kazakhstan, Kuwait, Morocco, Oman, Poland, Qatar, Saudi Arabia, Syria, Tunisia, United Arab Emirates, Uzbekistan, Vietnam.

Figure 2: Export concentration between resource and non-resource sectors with increasing oil dependency: all countries on the same boat.



Notes: concentration in resource exports, in all regime types. X-axis is oil dependency measures by oil rent share in GDP, data from the World Bank. Y-axis is the relative non-resource to resource exports from the WITS dataset. Panel (a) includes all countries in our dataset. Countries in panel (b) are: Australia, Austria, Canada, Columbia, Czech Republic, Denmark, France, Georgia, Germany, Greece, India, Israel, Italy, Japan, Lithuania, Malaysia, Netherlands, Norway, Russian Federation, Slovak Republic, Slovenia, South Africa, Sweden, Trinidad and Tobago, Turkey, Ukraine, United Kingdom, United States of America, Venezuela. Countries in panel (c) are: Algeria, Angola, Azerbaijan, Bahrain, Belarus, Cameroon, China, Congo Rep., Cote d'Ivoire, Cuba, Egypt, Gabon, Iran, Iraq, Jordan, Kazakhstan, Kuwait, Libya, Morocco, Oman, Qatar, Saudi Arabia, Syria, Tunisia, Turkmenistan, United Arab Emirates, Uzbekistan, Vietnam.

Figure 3: Number of one or more giant oilfield discoveries (from 1962 to 2003), by year

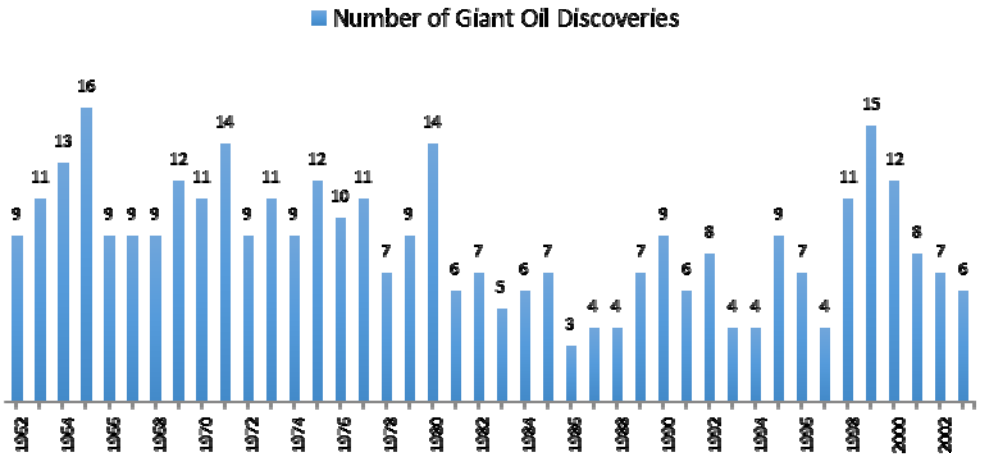
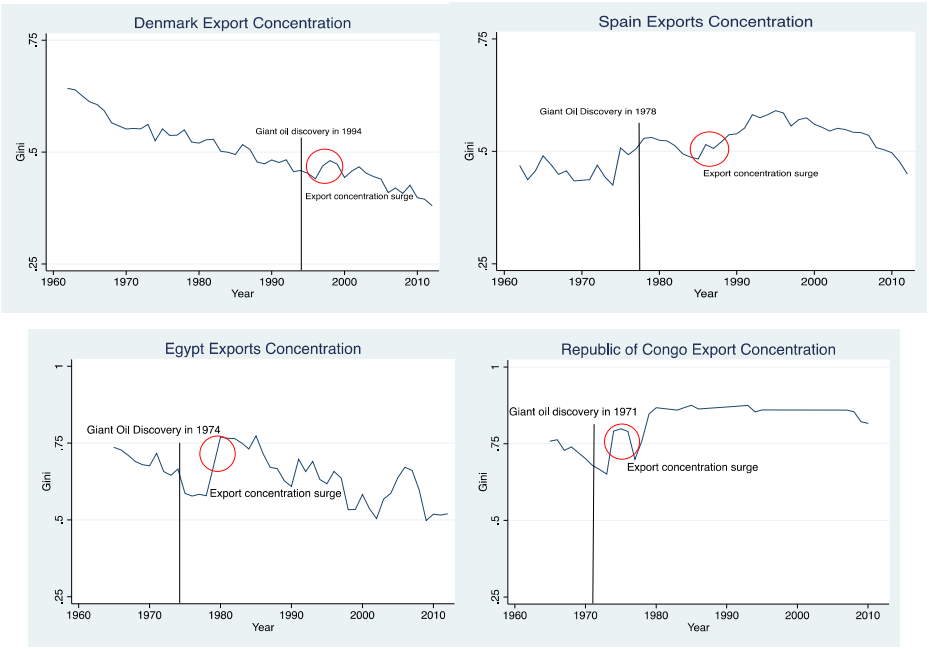
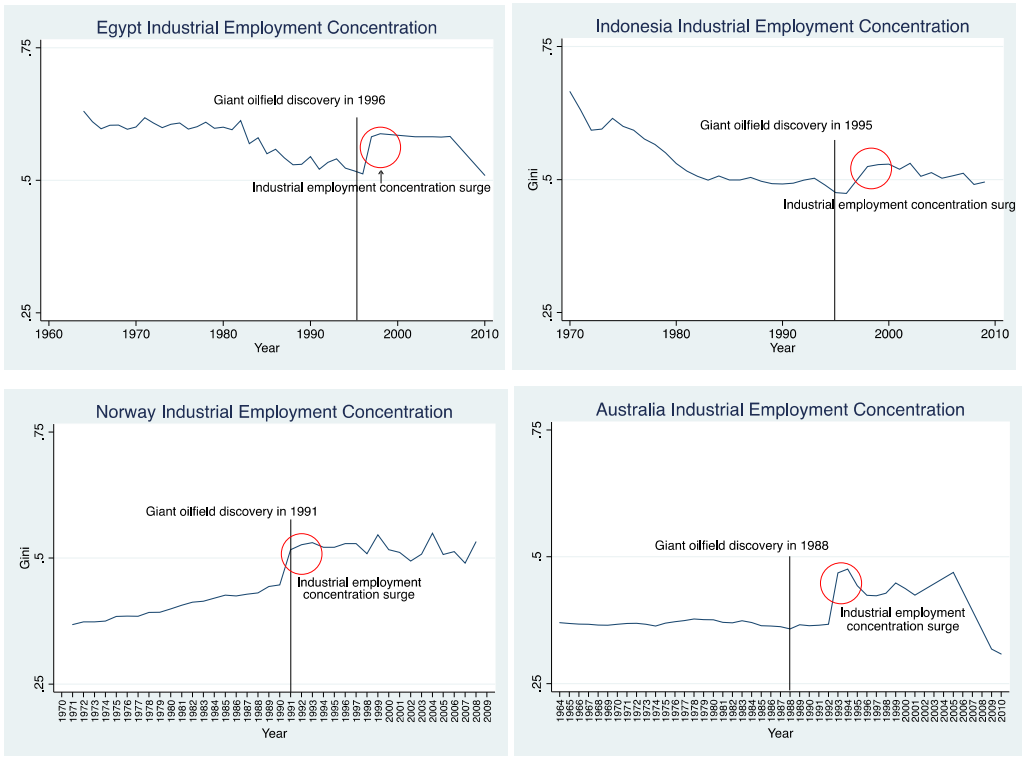


Figure 4: Oil discoveries and diversification in exports



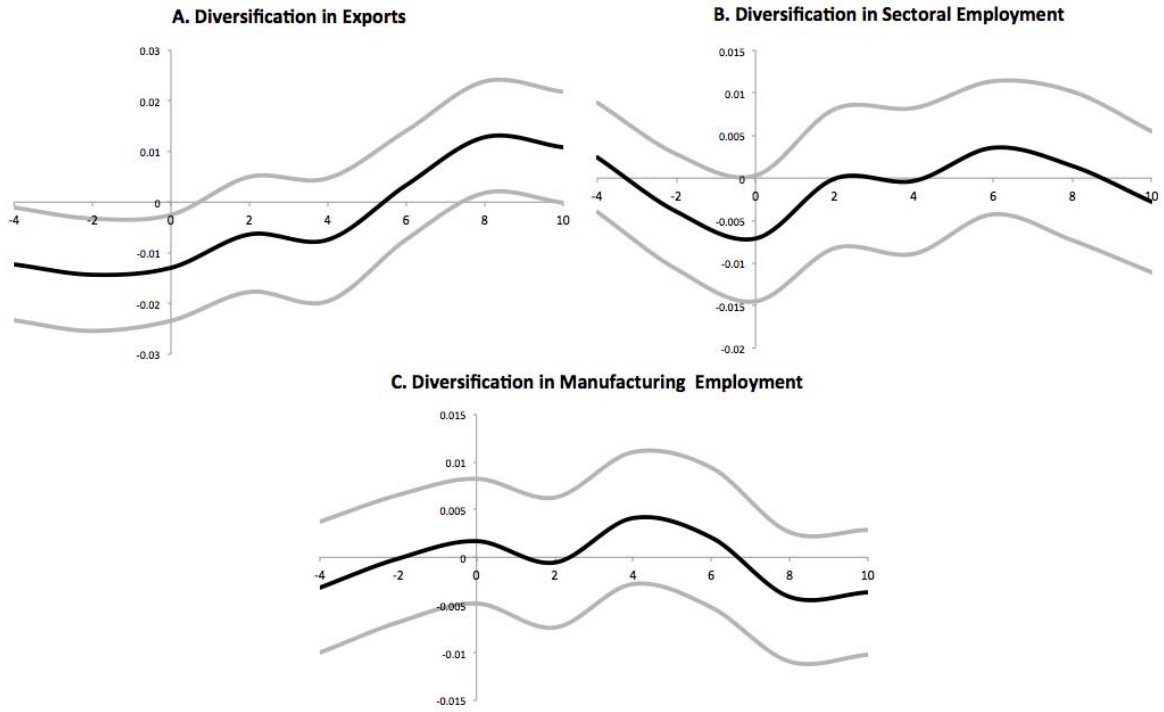
Notes: y-axis shows the Gini coefficient in each country, the x-axis shows the years where data is available. Gini ranges between 0 and 1, lower Gini indicates higher diversification. The vertical line shows the year of giant oil discovery in each country; the red circle shows the export concentration surge occurring after a giant oil discovery. *Data sources:* Exports data is from WITS. Oil discovery data is from Lei and Michaels (2014).

Figure 5: Oil discoveries and diversification in manufacturing employment



Notes: y-axis shows the Gini coefficient in each country, the x-axis shows the years where data is available. Gini ranges between 0 and 1, lower Gini indicates higher diversification. The vertical line shows the year of giant oil discovery in each country; the red circle shows the employment concentration surge occurring after a giant oil discovery. *Data sources:* manufacturing employment is from UNIDO. Oil discovery data is from Lei and Michaels (2014).

Figure 6: Giant Petroleum Discoveries and Economic Diversification



Notes: The x-axes report the number of years before or after t, ranging from t-4 to t+10. The black lines show the estimated coefficients and the gray lines show the 95% confidence intervals based on robust standard errors, which are clustered by country. All regressions control for previous discoveries (t-1 to t-10) and include country and year fixed effects. Details on variable construction can be found in the data section of the paper.

Table 1: summary statistics for the sectoral concentration indices

Variable	Obs	Mean	Standard Deviation (Overall)	Standard Deviation (between countries)	Standard Deviation (within countries)	Min.	Max.
ILO Employment (all sectors)							
Gini	2369	0.5028	0.0787	0.0919	0.0374	0.2540	0.8329
Theil Index	2369	0.4971	0.2230	0.2464	0.1360	0.1044	2.5860
HHI	2369	0.2273	0.0753	0.1004	0.0348	0.1562	0.9999
ILO Employment (non-resource sectors)							
Gini	2369	0.4524	0.0877	0.1023	0.0413	0.2540	0.8132
Theil Index	2369	0.4002	0.2094	0.2409	0.1175	0.1044	2.0630
HHI	2368	0.2307	0.0751	0.1011	0.0337	0.1590	0.8136
WITS Exports Diversification (all sectors)							
Gini	4577	0.6531	0.1286	0.1168	0.0652	0.3132	0.9
Theil Index	4576	0.9828	0.8018	0.6537	0.4968	0.1731	23.025
HHI	4554	0.3683	0.2059	0.1904	0.0950	0.1327	1
WITS Exports Diversification (non-resource sectors)							
Gini	4575	0.6243	0.1139	0.0997	0.0658	0.3077	0.8888
Theil Index	4574	0.8708	0.9329	0.6555	0.6931	0.1631	19.775
HHI	4558	0.3440	0.1590	0.1388	0.0901	0.1435	1
UNIDO Manufacturing Employment (employment)							
Gini	3564	0.5087	0.1086	.1109	.0435	0.2886	0.8823
Theil Index	3564	0.5313	0.3302	.4064	.1397	0.1482	3.0334
HHI	3558	0.1345	0.0850	.1016	.0280	0.0612	0.8742
Other Variables							
Oil discoveries	8933	0.0499	0.2178	0.1159	0.1843	0	1

Table 2: CORRELATION MATRICES FOR THE SECTORAL CONCENTRATION INDICES

	Gini	Theil Index	HHI
ILO Employment (all sectors)			
Gini	1.000		
Theil Index	0.897	1.000	
HHI	0.906	0.853	1.000
ILO Employment (non-resource sectors)			
Gini	1.000		
Theil Index	0.932	1.000	
HHI	0.926	0.917	1.000
WITS Exports Diversification (all sectors)			
Gini	1.000		
Theil Index	0.741	1.000	
HHI	0.897	0.802	1.000
WITS Exports Diversification (non-Resource sectors)			
Gini	1.000		
Theil Index	0.677	1.000	
HHI	0.894	0.745	1.000
UNIDO Manufacturing Employment			
Gini	1.000		
Theil Index	0.906	1.000	
HHI	0.727	0.803	1.000

Table 3: Number of years (from 1962 to 2003) with one or more giant oilfield discoveries, by country (treatment countries)

Country	Years	Country	Years	Country	Years
Former USSR	29	India	5	Albania	1
Iran	24	Algeria	4	Azerbaijan	1
Saudi Arabia	24	Argentina	4	Bangladesh	1
Australia	18	Colombia	4	Cote d'Ivoire	1
Nigeria	17	Congo, Rep.	4	Denmark	1
China	16	Kuwait	4	Ecuador	1
United States	16	Qatar	4	Equatorial Guinea	1
Norway	15	Peru	3	France	1
Indonesia	14	Thailand	3	Gabon	1
Brazil	13	Tunisia	3	Germany	1
United Arab Emirates	12	Bolivia	2	Hungary	1
United Kingdom	12	Brunei Darussalam	2	Morocco	1
Iraq	11	Italy	2	Namibia	1
Libya	11	Kazakhstan	2	New Zealand	1
Mexico	10	Myanmar	2	Papua New Guinea	1
Egypt, Arab Rep.	8	Netherlands	2	Philippines	1
Oman	8	Pakistan	2	Romania	1
Angola	7	Sudan	2	Russia	1
Canada	7	Trinidad & Tobago	2	Spain	1
Malaysia	6	Vietnam	2	Turkmenistan	1
Venezuela	6	Yemen	2		

Table 4: Do political and economic variables predict giant oil discoveries?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Previous year's polity2 score	0.005 (0.020)							
Previous year's sectoral employment diversification (Gini)		-0.593 (4.017)						
Previous year's manufacturing employment diversification (Gini)			-0.267 (2.278)					
Previous year's growth				-3.58e-14 (9.60e-14)				
Change in income pc					-0.000064 (0.00012)			
Change in government expenditure						-0.01744 (0.2289)	-0.01186 (0.0095)	
Change in investments						0.03596 (0.0229)		0.02772 (0.02051)
Observations	2672	772	1437	2092	2256	481	1057	481
Number of countries	111	67	91	78	104	48	76	47
Years	1952-2003	1971-2003	1965-2003	1963-2003	1953-2003	1983-2003	1963-2002	1983-2003

Notes: reported coefficients are from a fixed-effects logit model of the probability of a giant oil discovery occurring in a given year. Robust standard errors clustered at the country level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Giant Oil Discovery and Diversification

Outcome in year:	t+2	t+4	t+6	t+8	t+10
Panel A. Diversification in Non-Resource Exports					
Discovery	-0.006	-0.007	0.003	0.013**	0.011*
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	3677	3889	3971	3936	3900
R^2	0.78	0.77	0.77	0.77	0.77
No. of countries	126	126	126	126	124
Years covered	1963-2005	1963-2007	1963-2009	1963-2011	1963-2012
Panel B. Diversification in ILO sectoral employment					
Discovery	-0.0007	-0.0003	0.003	0.001	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	2049	2191	2232	2205	2178
R^2	0.84	0.826	0.817	0.814	0.813
No. of countries	111	112	112	111	111
Years covered	1970-2005	1970-2007	1970-2008	1970-2008	1970-2008
Panel C. Diversification in UNIDO manufacturing employment					
Discovery	-0.0005	0.004	0.002	-0.004	-0.0036
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	3120	3244	3289	3263	3235
R^2	0.871	0.868	0.867	0.866	0.866
No. of countries	120	119	119	119	119
Years covered	1964-2005	1964-2007	1964-2008	1964-2008	1964-2008

Notes: Gini index is the dependent variable. Past Discoveries: the number of years with discoveries from t-10 to t-1. Data sources: (A) exports data is from WITS. (B) Sectoral employment is from ILO, (C) manufacturing employment is from UNIDO. Robust standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Non-giant Oil Discovery and Diversification

Outcome in year:	t+2	t+4	t+6	t+8	t+10
Panel A. Non-Resource Exports					
Discovery	-0.006	-0.007	0.003	0.013**	0.011*
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	3677	3889	3971	3936	3900
No. of countries	57	57	57	57	57
Years	1963-2005	1963-2007	1963-2009	1963-2011	1963-2012
Panel B. Non-Resource Sectoral employment					
Discovery	-0.0001	-0.0001	0.004	0.001	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	2049	2191	2232	2205	2178
No. of countries	55	55	55	54	54
Years	1970-2005	1970-2007	1970-2008	1970-2008	1970-2008
Panel C. Manufacturing employment					
Discovery	-0.001	0.004	0.002	-0.004	-0.004
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	3120	3244	3289	3263	3235
No. of countries	55	55	55	55	55
Years	1964-2005	1964-2007	1964-2009	1964-2010	1964-2010

Notes: Gini index is the dependent variable. Past Discoveries: the number of years with discoveries from t-10 to t-1. Data sources: (A) exports data is from WITS. (B) Sectoral employment is from ILO, (C) manufacturing employment is from UNIDO. Robust standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Giant Oil Discovery Size and Diversification

Outcome in year:	t+2	t+4	t+6	t+8	t+10
	Panel A. Discovery size in quartile 1				
Non-Resource Exports	-0.007	-0.013	-0.003	0.002	0.017**
	(0.011)	(0.012)	(0.010)	(0.009)	(0.009)
No. of countries	126	126	126	126	125
Years	1963-2005	1963-2007	1963-2009	1963-2011	1963-2012
Sectoral Employment	0.004	0.007	0.009	-0.002	-0.002
	(0.007)	(0.006)	(0.058)	(0.008)	(0.007)
No. of countries	111	112	112	111	111
Years	1970-2005	1970-2007	1970-2008	1970-2008	1970-2008
Manufacturing Emp.	0.002	0.004	0.001	-0.001	-0.007
	(0.005)	(0.005)	(0.005)	(0.004)	(0.006)
No. of countries	120	119	119	119	119
Years	1964-2005	1964-2007	1964-2009	1964-2010	1964-2010
	Panel B. Discovery size in quartile 2				
Non-Resource Exports	-0.004	-0.002	0.002	0.010	0.006
	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)
No. of countries	126	126	126	126	125
Years	1963-2005	1963-2007	1963-2009	1963-2011	1963-2012
Sectoral Employment	-0.004	-0.005	0.004	0.013**	-0.001
	(0.006)	(0.007)	(0.006)	(0.006)	(0.005)
No. of countries	111	112	112	111	111
Years	1970-2005	1970-2007	1970-2008	1970-2008	1970-2008
Manufacturing Emp.	-0.009	-0.004	-0.002	-0.007	-0.008
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
No. of countries	120	119	119	119	119
Years	1964-2005	1964-2007	1964-2009	1964-2010	1964-2010
	Panel C. Discovery size in quartile 3				
Non-Resource Exports	0.001	-0.000	0.004	0.006	0.003
	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
No. of countries	126	126	126	126	125
Years	1963-2005	1963-2007	1963-2009	1963-2011	1963-2012
Sectoral Employment	-0.005	-0.005	0.001	-0.004	-0.001
	(0.007)	(0.009)	(0.007)	(0.008)	(0.007)
No. of countries	111	112	112	111	111
Years	1970-2005	1970-2007	1970-2008	1970-2008	1970-2008
Manufacturing Emp.	0.003	0.008	0.004	-0.006	0.005
	(0.007)	(0.006)	(0.007)	(0.006)	(0.005)
No. of countries	120	119	119	119	119
Years	1964-2005	1964-2007	1964-2009	1964-2010	1964-2010
	Panel D. Discovery size in quartile 4				
Non-Resource Exports	-0.003	0.005	0.012	0.021***	0.011
	(0.012)	(0.010)	(0.009)	(0.009)	(0.009)
No. of countries	126	126	126	126	125
Years	1963-2005	1963-2007	1963-2009	1963-2011	1963-2012
Sectoral Employment	0.004	-0.009	-0.009	-0.006	-0.005
	(0.008)	(0.012)	(0.011)	(0.007)	(0.007)
No. of countries	111	112	112	111	111
Years	1970-2005	1970-2007	1970-2008	1970-2008	1970-2008
Manufacturing Emp.	0.008	0.010	0.009	0.009	0.005
	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)
No. of countries	120	119	119	119	119
Years	1964-2005	1964-2007	1964-2009	1964-2010	1964-2010

Notes: Gini index is the dependent variable. Discovery size normalized by population. Robust standard errors clustered at the country level in parentheses. All regressions include previous discoveries over the past ten years, country and year fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Oil Discovery, Political Institutions and Diversification

Outcome in year:	t+2	t+4	t+6	t+8	t+10
Panel A. Non-Resource Exports					
Discovery	-0.005	-0.007	0.005	0.013**	0.011*
	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)
Discovery*Polity2(t-1)	-0.001	0.002	-0.007	-0.001*	-0.008
	(0.006)	(0.007)	(0.006)	(0.0006)	(0.006)
Polity2(t-1)	0.002	-0.002	-0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	3500	3703	3781	3746	3710
R^2	0.77	0.77	0.77	0.77	0.77
No. of countries	122	122	122	122	122
Years	1963-2005	1963-2007	1963-2008	1963-2008	1963-2008
Panel B. Non-Resource Sectoral Employment					
Discovery	0.01	0.01	0.07	0.03	0.04
	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
Discovery*Polity2(t-1)	-0.003	-0.004	-0.008	-0.005	-0.01**
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Polity2(t-1)	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	1981	2119	2158	2131	2104
R^2	0.858	0.843	0.835	0.833	0.832
No. of countries	107	108	108	108	107
Years	1970-2005	1970-2007	1970-2008	1970-2008	1970-2008
Panel C. Manufacturing Employment					
Discovery	0.01	0.06	0.03	-0.05	-0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Discovery*Polity2(t-1)	-0.01**	-0.01***	-0.01**	-0.001	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Polity2(t-1)	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	2955	3077	3121	3095	3067
R^2	0.878	0.875	0.873	0.873	0.873
No. of countries	115	115	115	115	115
Years	1964-2005	1964-2007	1964-2008	1964-2008	1964-2008

Notes: Gini index is the dependent variable. Past Discoveries: the number of years with discoveries from t-10 to t-1. Data sources: (A) exports data is from WITS. (B) Sectoral employment is from ILO, (C) manufacturing employment is from UNIDO. Robust standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Oil discovery, Executive Constraints and Diversification

Outcome in year:	t+2	t+4	t+6	t+8	t+10
Panel A. Non-Resource Exports					
Discovery	-0.004	-0.005	0.001	0.020***	0.012**
	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)
Discovery*xconst (t-1)	-0.004	-0.004	0.005	-0.008**	-0.005
	(0.005)	(0.004)	(0.005)	(0.003)	(0.004)
Executive constraints	0.007	0.003	0.002	0.004	0.002
(t-1)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	3516	3723	3803	3768	3732
R^2	0.77	0.77	0.77	0.77	0.77
No. of countries	123	123	123	123	121
Years	1963-2005	1963-2007	1963-2008	1963-2008	1963-2008
Panel B. Non-Resource Sectoral Employment					
Discovery	-0.003	-0.0002	0.002**	0.002	0.002**
	(0.004)	(0.004)	(0.001)	(0.012)	(0.001)
Discovery*xconst(t-1)	0.005**	-0.0001	-0.004**	-0.003*	-0.005***
	(0.0002)	(0.0002)	(0.001)	(0.0019)	(0.0018)
Executive constraints	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
(t-1)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	1981	2119	2158	2131	2104
R^2	0.85	0.83	0.82	0.82	0.82
No. of countries	107	108	108	107	107
Years	1970-2005	1970-2007	1970-2008	1970-2008	1970-2008
Panel C. Manufacturing Employment					
Discovery	-0.001	0.004	0.002	-0.004	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Discovery*xconst(t-1)	0.002	-0.0001	0.004*	-0.001	-0.002
	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)
Executive constraints	0.001*	0.001*	0.0008	0.001*	0.001*
(t-1)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0007)
Controls	Past Discoveries, Country Dummies, Year Dummies				
Observations	2959	3081	3125	3099	3071
R^2	0.88	0.88	0.87	0.87	0.87
No. of countries	116	115	115	115	115
Years	1964-2005	1964-2007	1964-2008	1964-2008	1964-2008

Notes: Gini index is the dependent variable. Past Discoveries: the number of years with discoveries from t-10 to t-1. Data sources: (A) exports data is from WITS. (B) Sectoral employment is from ILO, (C) manufacturing employment is from UNIDO. Robust standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Oil Discovery, Political Institutions and Diversification: IV Approach

Outcome in year:	t+2	t+4	t+6	t+8	t+10
Panel A. Non-Resource Exports					
Discovery	-0.002	-0.001	-0.002	-0.001	0.003**
	(0.011)	(0.007)	(0.006)	(0.001)	(0.0008)
Discovery*Polity2(t-1)	0.006	0.003	0.001	0.007	0.004
	(0.005)	(0.003)	(0.003)	(0.006)	(0.004)
Controls	Past Discoveries, Country Dummies, Year Dummies, Polity2(t-1)				
Kleibergen-Paap F stat	7.71	4.40	0.16	10.11	5.19
Stock-yogo critical value	19.93/7.25	19.93/7.25	19.93/7.25	19.93/7.25	19.93/7.25
Observations	3500	3703	3781	3746	3710
R^2	0.77	0.77	0.77	0.77	0.77
Panel B. Non-Resource Sectoral Employment					
Discovery	0.012	0.002	-0.035	0.015**	-0.003
	(0.008)	(0.008)	(0.047)	(0.007)	(0.002)
Discovery*Polity2(t-1)	-0.002	-0.003	0.008	-0.003**	0.005
	(0.002)	(0.011)	(0.011)	(0.001)	(0.004)
Controls	Past Discoveries, Country Dummies, Year Dummies, Polity2(t-1)				
Kleibergen-Paap F stat	1.72	6.013	0.25	2.39	3.66
Stock-yogo critical value	19.93/7.25	19.93/7.25	19.93/7.25	19.93/7.25	19.93/7.25
Observations	1981	2119	2158	2131	2104
R^2	0.858	0.843	0.835	0.833	0.832
Panel C. Manufacturing Employment					
Discovery	0.007*	0.002	0.006*	0.005***	0.009
	(0.0035)	(0.002)	(0.0037)	(0.0018)	(0.007)
Discovery*Polity2(t-1)	-0.004**	-0.011	-0.0032*	-0.003***	-0.006
	(0.001)	(0.010)	(0.002)	(0.001)	(0.004)
Controls	Past Discoveries, Country Dummies, Year Dummies, Polity2(t-1)				
Kleibergen-Paap F stat	2.99	1.0	2.04	9.25	1.46
Stock-yogo critical value	19.93/7.25	19.93/7.25	19.93/7.25	19.93/7.25	19.93/7.25
Observations	2955	3077	3121	3095	3067
R^2	0.878	0.875	0.873	0.873	0.873

Notes: Gini index is the dependent variable. Past Discoveries: the number of years with discoveries from t-10 to t-1. Instrumental variables are the log (out-of region natural disaster), and the log (oil reserves per capita) and their interaction with Polity2 for instrumenting the interaction term (Discovery*Polity2). Data sources: (A) exports data is from WITS. (B) Sectoral employment is from ILO, (C) manufacturing employment is from UNIDO, instruments from Cotet and Tsui (2013). Robust standard errors clustered at the country level in parentheses. All coefficients (and standard errors) are multiplied by 1000 to improve readability. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$